David Maddock

Dr. Michael Drout

LITZ 6399: Master's Thesis

April 23, 2018

The composite nature of *Andreas*

Introduction

Cynewulf has received scrutiny as one of the few named poets in the
Anglo-Saxon vernacular due to the existence of four poems--*Elene, Juliana, Christ II,*
and *The Fates of the Apostles*--that contain rune signatures bearing his name. Much ink
has been spilled in analysis of these as well as related unsigned poems deemed to
have possible Cynewulf authorship. Because we know so little about Cynewulf's
biography, "the whole question of Cynewulf's authorship is intimately tied up with the
issue of style" ("Style" 271). Foremost among the broader "Cynewulfian group" of
poems sharing stylistic resemblance to the signed poems is *Andreas*. Scholars in the
early nineteenth century were inclined to attribute the poem to Cynewulf on this account
(Bjork xvi). However, later scholars have argued against this possibility based on
grammatical and metrical differences as well as its shared diction with *Beowulf* (Fulk
7-8; Greenfield 165). Scholars on both sides of the authorship question have deployed
stylistic data in their arguments for and against *Andreas*, but interrogating this data for
what it might tell us about the structure and composition of the poem itself has not been

of primary concern. Before deploying such an argument, the critic must demonstrate the stylistic uniformity of the work in question since it would be inappropriate to generate a single authorial profile by conflating the contributions of multiple authors; however, such a demonstration has not been made with *Andreas*.

The practice of repurposing poetic material is not without precedent in the Old English poetic corpus. Within the Cynewulfian group, it is now taken for granted that the *Christ* and *Guthlac* poems found in the Exeter book are comprised of five independent poems rather than two long works (Fulk 5-9). Several other poems have also been shown to have a composite nature. In 1875 Eduard Sievers noted a stylistic division within the poem *Genesis*, applying philological principles to identify a section of lines (now called *Genesis B*) which were translated from Old Saxon (Greenfield 209-10). Drout and Chauvet use lexomic techniques to show that the "Song of the Three Youths" poem embedded within *Daniel* is derived from an Old English version rather than Latin as supposed by Paul Remley (298-9). Most recently, Leonard Neidorf has marshalled compelling metrical and lexical evidence to demonstrate that *The Dream of the Rood*, a poem from the same manuscript as *Andreas*, can be divided into two components. The author of the first 77 lines exhibits a superior command of Old English poetics to that of the latter half, who may have been motivated to restore a damaged copy of the poem (68-70).

This paper re-examines the distribution of stylistic features within *Andreas* to determine to what extent they are uniform, and it employs statistical analysis to describe and explain any observed variance. Such an investigation is a necessary prerequisite

for any endeavor dependent on identifying an author's style. The discovery of unusual distributions of differentiated markers of style can be a hallmark of a composite work. "If sections of a work exhibit linguistic distinctions that cannot reasonably be attributed to chance or to the deliberation of a single author, credence in a theory of composite authorship appears warranted" (Neidorf 53). Identifying a poem's composite nature may also shed new light on dating the work, locate its place of composition, provide new historical information about editorial and poetic composition practices, and recover the outline of an earlier lost work.

Methodology

Scholars such as George Krapp have compiled lists of similarities and differences in terms of grammar and diction between *Andreas* and the signed poems. More recently, Andy Orchard has identified distinctive formulaic parallels and word compounds among these as well as between *Andreas* and *Beowulf*. However, neither Krapp nor Orchard considered whether such stylistic elements are uniformly or heterogeneously distributed within the poem, and both scholars neglected to include a visual component which, as Franco Moretti puts it, "shows us that there is something that needs to be explained" (39). By plotting such data as a function of line number, it is possible to visualize and interpret the distribution of specific elements. Having done so, the data can be further correlated with other quantitative techniques such as rolling

window analysis and hierarchical, agglomerative cluster analysis as applied to Anglo-Saxon studies by the Lexomics Research Group at Wheaton College.[1]

Whereas the oral-formulaic analysis works on the level of phrases, cluster analysis quantifies the frequency of individual words. Cluster analysis generates a statistical profile of an author's style by measuring the frequency with which an author uses each word in a sample. Similar techniques--such as John Burrow's Delta procedure--have been used to investigate questions of authorship in other domains, but these methods presuppose the critic has access to a corpus of known work from a "closed" field of possible authorial candidates ("Questions" 8-10; Hoover). The Delta procedure collapses an author's statistical profile into a single "delta-score" defined as "the mean of the absolute differences between z-scores for a set of word-variables" ("Delta" 271). Burrows limited his analysis to the top 150 words in order to avoid the effects of domain-specific content words (469).[2] Cluster analysis replaces the delta-score calculation with dendrogramming, the visual representation of relative affinity between texts, which allows word frequency profiles to be applied more efficaciously to "open" fields of largely unknown authors (eg. the Old English poetic corpus) as well as within a single work to investigate questions of internal structure (*Unlocked* 7-8; "Dendrogrammatology" 305). Since cluster analysis does not use z-scores, the contribution of content words to the difference calculation is attenuated

---

[1] Moretti also notes that quantitative analysis presupposes a formalism that "makes quantification possible in the first place" (25). As such, it is important that the feature selected for quantification be grounded in a solid theoretical basis. The features quantified here--oral-formulaic phrases, word choice, and orthographic preferences--have long been cited by scholars as aspects of literary style.

[2] The use of z-scores ensures that each word is considered of equal weight despite the fact that word frequency declines geometrically according to Zipf's law (Pierce 294). Hoover demonstrated that the Delta procedure can be extended to the top 600-700 most frequent words with larger sample sizes.

naturally by the effect of Zipf's law and the fact that every text is directly compared to every other text, rather than against an average baseline.

One limitation of cluster analysis, especially when applied to questions of internal structure, is that it requires texts to be divided into "chunks" that are compared relative to each other. The size and placement of the divisions can affect outcomes of the analysis, such as muting the effect of a particular feature or creating chunks that are too small to measure word frequency with sufficient resolution (*Unlocked* 11-2).[3] Rolling window analysis is a complementary technique which overcomes this problem by generating a running calculation within a continuously shifting window of fixed size ("Tracking" 291). The sensitivity of the indicator is adjusted by changing the size of the window used. Grammatical, stylistic, metrical, or orthographic features of a text can be analyzed in this way. Drout and Chauvet have applied this technique to Old English texts by calculating a moving ratio of þ to ð. Because the two characters were often used interchangeably, changes in the trend of this ratio can correlate with differences in source, author, or scribal hand (315-6). Since rolling window analysis can be applied to orthographic data, it can be used to extract a line of evidence independent of word choice, syntax, and morphology.

As each methodological approach uses a different quantitative basis--the phrase, the word, the symbol--correlation between them can be a strong indication that a finding is not an artifact from a particular technique. Each of these methods present data in a

---

[3] Zipf's law is again the enemy. A chunk of size n cannot measure accurately the frequency of words which occur with probability less than $1/n$. As chunk size increases so does the resolution available for measuring word frequency. This enables the frequency of more words to be measured and therefore the percentage of the chunk quantified increases.

way that allows "'emerging' qualities, which were not visible at the lower level" to be discovered (Moretti 53). When used in conjunction with more traditional qualitative approaches, quantitative tools provide new and complementary methods of interpretation that could lead to fresh insights.

**The distribution of formulaic language in *Andreas***

In his summary of the state of Cynewulf criticism, Fulk notes that "parallel passages [...] carry little weight now that oral-formulaic theory has shown the pervasiveness of formulae and their public, conventional nature" (5). However, Orchard asserts that "the 'white noise' of traditional and inherited formulae" can be filtered out by limiting one's scope to specific formulaic language unique to the works under inquiry ("Style" 273). One must refrain from absolute pronouncements for or against the usefulness of formulaic language *per se*. Some formulae considered distinctive today by Orchard's system were probably widespread at the time while others only seem commonplace now due to accidents of history. Nevertheless, if certain formulae were unevenly distributed throughout a literary or oral tradition by region, time period, or "school" then Orchard's approach should capture these tendencies in the aggregate despite uncertainty regarding any particular formula on its own. Furthermore, the efficacy of this approach can be confirmed when correlated with other independent methods.

Having quantified a large number of shared formulae between *Andreas* and the signed poems, Orchard concludes that his "figures strongly suggest either unity of

authorship or conscious literary borrowing," opting for the latter on the basis of Fulk's evidence (which extended Krapp's original argument) of divergent diction (287). However, Orchard's data can be applied to more than just the authorship question.[4] When plotted by line number, it becomes apparent that Cynewulfian parallels are not evenly distributed throughout the poem (Figure 1). The graph[5] shows four sections of *Andreas* where no such parallels have been found. It is important to note that this does not mean that formulaic language is wholly absent from these sections, but rather that distinctively Cynewulfian formulae are missing.
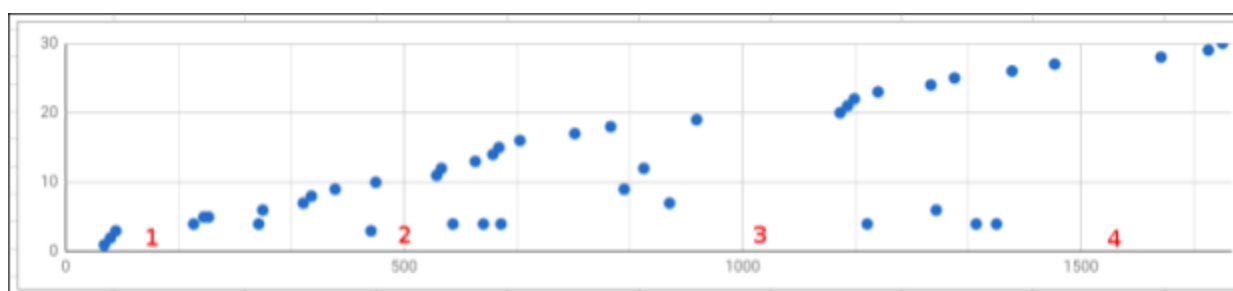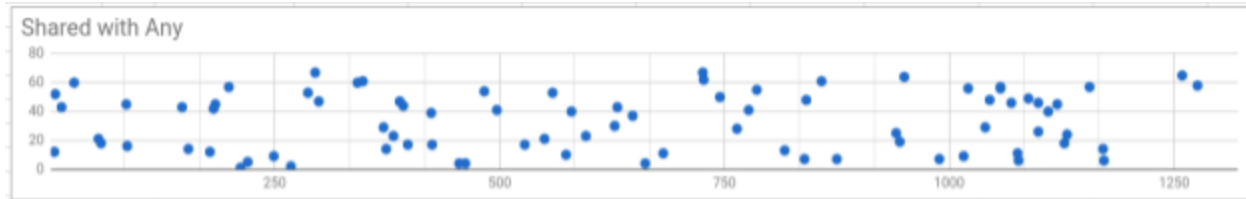

Figure 1: The distribution of Cynewulfian formulae in *Andreas*

As a comparison, the following graph plots the Cynewulfian formulae from *Elene*. *Elene* is an ideal control for comparison as it is signed by Cynewulf, of similar length, and from the same manuscript as *Andreas*. With the exception of the very end of the poem, Cynewulfian formulae are evenly distributed without significant gaps.

---

[4] There remains a plethora of formulaic data compiled within the Cynewulfian group which can be analyzed as this paper describes which might shed more light on the nuances of Cynewulf's style, his methods of composition, and how the rest of the group may relate to Cynewulf and each other.

[5] The vertical scale represents an arbitrary identification number assigned to each formula by Orchard, so the linearly increasing pattern of the *Andreas* plot is not significant. The horizontal distribution is what is of concern here.

Figure 2: The distribution of Cynewulfian formulae in *Elene*

Scholars generally agree that Cynewulf used a Latin prose exemplar when composing *Elene* and yet *Elene* does not exhibit a similar pattern in its formula distribution where Latin source influence is detected (Anderson 103). Likewise, an analysis of where *Andreas* differs from its closest Latin exemplar in the Casanatensis manuscript does not show any clear correlation with the discontinuities in Cynewulfian formulaic distribution (Friesen 301-7). This difference in distribution pattern suggests that these sections lacking Cynewulfian formulae in *Andreas* have a distinctly non-Cynewulfian bias. The fact that the formulaic language is discontinuous rather than relatively less frequent suggests that the source was not merely a reference--such as a Latin prose narrative like the *Vita Cyriaci* or the Casanatensis prose exemplar--but an Old English poem which could be incorporated into a larger poem with little editorial work.

The narrative content of *Andreas A* and *B*

For the purpose of analysis, the four sections that show no distinctive Cynewulfian formulae as described above are called *Andreas B* while the parts in between B sections are called *Andreas A*. This divides the poem into nine parts as outlined in Table 1. Six of these breaks appear in the middle of a sentence or thought,

but the boundaries have not been modified to coincide with syntactical hints so as to preserve the quantitative basis for the divisions.

Table 1: A/B divisions in *Andreas*

	Start line	End line	Summary
Intro	1	74	The Mermedonian situation and how Matthew came to be imprisoned is explained.
B1	75	188	Matthew prays for deliverance. God responds, then calls Andrew to rescue him.
A1	189	458	Andrew responds to God skeptically and is rebuked. Andrew finds a boat piloted by God in disguise. After discussing terms of payment, they set off only to encounter a storm. Andrew calms his men by explaining God's power over weather.
B2	459	547	The storm abates. Andrew complements the skill of the captain. The captain affirms God's power over the weather, attributing the storm's passing to Andrew's piety. Andrew prays in thanksgiving.
A2-3[6]	548	932	The captain asks Andrew why the Jews did not recognize Jesus as the Messiah; Andrew responds. They discuss the miracles of Jesus including the story of a stone angel brought to life and the resurrection of patriarchs, who proclaimed Jesus yet their hearers did not believe. Andrew sleeps. Upon waking, Andrew and his crew find themselves in Mermedonia and realize they had encountered the divine. Christ appears again, reminding Andrew of his earlier skepticism.
B3	933	1143	Christ reminds Andrew of his mission and explains Andrew's coming suffering. Christ departs. Andrew rescues Matthew and frees the prisoners, killing several guards in the process.[7] Matthew and the prisoners escape the city while Andrew remains. Discovering the empty

---

[6] The sub-division of this section into A2 and A3 is explained later.

[7] Note there are lost pages in the manuscript here. Stevick estimates the lacuna comprises 78 missing lines (109). The assumption that the lost lines would be contiguous with B appears reasonable but is, of course, speculation.

			prison, the Mermedons eat the corpses of the guards. Casting lots for the next meal, the lot falls to a general who gives his innocent son instead. Andrew is outraged.
A4	1144	1461	God protects the boy; famine ensues. The devil appears and incites the crowd to find and eat Andrew. They find Andrew, bind and imprison him. He is tortured for three days but no one can kill him because of the sign of the cross on his forehead. Andrew prays for death. God responds, showing how flowers have grown where Andrew's blood was spilled.
B4	1462	1617	God restores Andrew's body. The poet interjects to express humility and reference "fyrnsægen" (old saying, ancient tradition). Andrew commands a flood to appear out of a stone pillar. Flood and fire kill many and convince the people of their wrongdoing. Andrew calms the flood. The worst heathens had been killed. The survivors recognize God's power.
End	1618	1722	The city converts to Christianity. A church is built on the site of the pillar. Andrew is induced to remain seven days before sailing home.

A close reading of the B sections shows that a coherent outline of Andrew's story is described therein. Taken on its own, it tells the story of Andrew's calling by God, the rescue of Matthew from the Mermedons, and the cleansing of the city with almost all of the key plot points included. It does so using only 33% of the extant lines. However, there are two notable plot holes in B. First, the reader is unsure how Andrew finds himself in the midst of a storm at sea without A1. Second, the actual capture of Andrew by the Mermedons is missing without A4. How these components of B may be recovered is discussed in the next section.

The A sections do not significantly advance the plot. These parts are preoccupied with events on Andrew's journey and his torture in prison. Despite the absence of major plot points, they comprise 56% of the extant lines. (As literal edge cases, the two pieces containing the introduction and epilogue are set aside for now.) Nevertheless, A has its own exclusive themes. The topic of Andrew's skepticism is treated only in A and frames stories of unbelief in the face of miraculous signs. The explicit identification of the helmsman as God is found almost entirely in A as is Andrew's torture in simulation of Christ's passion.

A adds several aspects to the sea voyage passage which change the whole character of the episode. For instance, the conflict in B's sea voyage is the storm, not the distance. There is no mention in B that a three-day voyage to Mermedonia from Achaia would require miraculous intervention; this is an A concern. Also, if not for the poet's use of "ece dryhten" (eternal Lord) in reference to the captain in line 510, B2 would read like an exchange between pious mortals which may or may not intend to imply the helmsman is God. A later editor wishing to make that implication explicit might include material such as that found in A.

The A sections also create problems that don't exist for B when it stands on its own. The famine and subsequent torture of Andrew make the poem less coherent. In B alone, the precise capture of Andrew is unclear--perhaps it is the result of an attempt to rescue the innocent boy--but there are no obvious inconsistencies present. Andrew "thought [the boy's plight] miserable" (*earmlic þuhte*), but takes no action as the poem crosses over to A where God's perfunctory rescue of the boy renders the whole episode

rather pointless (1135; 1143b-44). What's more, it sets off a period of famine predicated on the notion that cannibals which have been shown capable of eating their own kind no longer do so (1155-62a). During this time the reader must presume that Andrew is living voluntarily in a famine-wracked city when he could have left with Matthew or any time thereafter. And once captured, the starving citizens show little interest in actually eating him (1249-50a). From a narrative perspective, A4 can be interpreted as a justification to extend the story with a passion scene.

Nevertheless, these events are not the invention of the A poet as they also appear in the prose retelling of the Casanatensis manuscript (Friesen 301-7; Greenfield 159). It is likely then that the A poet knew or referenced a source like Casanatensis. Likewise, the absence of these features in B suggests that the B poet did not learn the story from such a written source. In fact, a desire to harmonize the B poem with a Latin prose exemplar may be the reason the project of revising B was taken up.

Both A and B sections are internally consistent but several plot problems manifest when they are read together. The B material contains the core of Matthew's rescue narrative and is characterized by an absence of distinctive Cynewulfian formulae. However, the A material constitutes asides or digressions of a homiletic or didactic nature and is highly correlated with such formulae. This dichotomy suggests that *Andreas B* represents a non-Cynewulfian source which has been expanded upon.

Affinities with *Beowulf* in *Andreas B*

In "The Originality of *Andreas*," Orchard compiles a list of parallel passages and distinctive compound words used between *Andreas* and *Beowulf*. A third dataset enumerates the uniquely shared compounds between *Andreas* and the signed Cynewulf poems. A fourth lists compounds unique to *Andreas* alone. By close investigation of these lists, Orchard argues that the *Andreas* poet innovates new compounds in dialogue with language borrowed from *Beowulf* and the signed poems (333). Table 2 tabulates these occurrences against the A/B sections described above. Each absolute count is corrected for the size of the section, resulting in an "occurrence per line" score. Those scores deviating more than 20% from the mean for its group have been flagged green (if above) or red (if below).

Table 2: Distribution of word compounds in *Andreas* by section

	Beowulf parallels	Per line	*Beowulf* compounds	Per line	Cynewulf compounds	Per line	Unique compounds	Per line
Intro	1	0.014	1	0.014	3	0.041	6	0.081
B1	2	0.018	3	0.026	2	0.018	22	0.193
A1	8	0.030	9	0.033	6	0.022	32	0.119
B2	2	0.022	4	0.045	1	0.011	16	0.180
A2-3	3	0.008	2	0.005	7	0.018	26	0.068
B3	3	0.014	8	0.038	6	0.028	26	0.123
A4	2	0.006	5	0.016	10	0.031	29	0.091
B4	2	0.013	3	0.019	3	0.019	23	0.147
End	0	0.000	2	0.019	2	0.019	7	0.067
	Mean	0.013	**Mean**	0.021	**Mean**	0.023	**Mean**	0.109
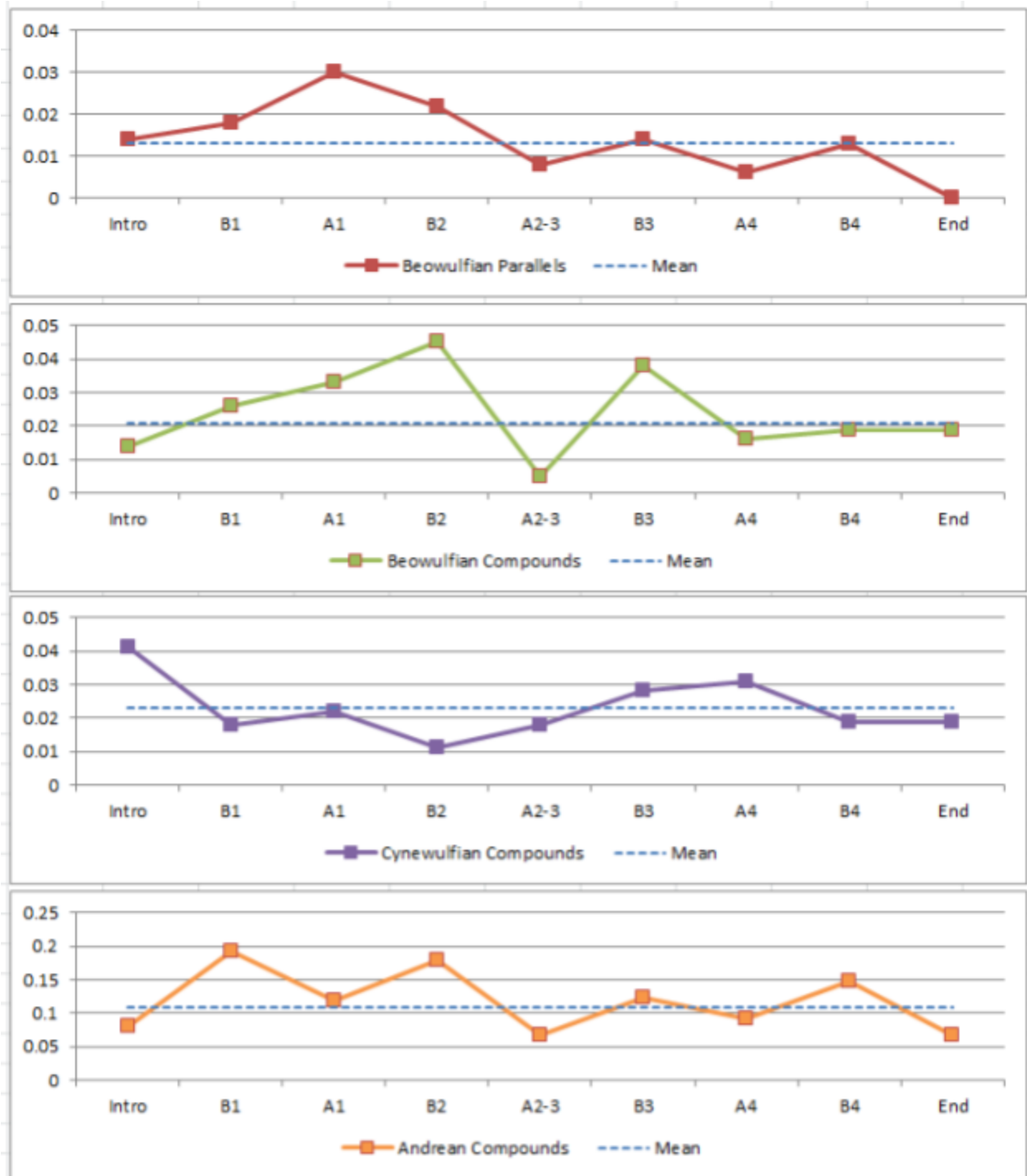
Figure 3: Distinctive compounds per line for each section

Figure 3 plots the "occurrence per line" scores from Table 2 against the mean for

each metric.

  The Beowulfian parallels and compound words are present particularly in the sea voyage episode (B1-A1-B2). B3 is also heavily weighted with Beowulfian compound words. With the exception of A1, A sections exhibit low affinity to *Beowulf* by this metric. Despite its high Beowulfian scores, A1 has the highest Cynewulfian affinity as measured by word compounds among the sea voyage triad. The Cynewulfian word compounds are distributed more broadly throughout the poem but three of the four B sections have scores below the mean--also supporting the interpretation that B has a non-Cynewulfian bias. Despite lacking distinct Cynewulfian formulae, B3 registers high on the Cynewulfian compounds metric. This indicates that this section has been reworked more by the A poet than the rest of the B material. The dataset intended to capture the most innovative word compounds--those compounds unique to *Andreas*--also disproportionately occur in the B sections (340). Nevertheless, innovative compounds are also well attested in A sections which implies that this phenomena is not original to the B material. Therefore, the originality in word coinage Orchard identified as a stylistic hallmark of *Andreas* is in part an editorial technique used by the A poet to retouch the Beowulfian sections.

  A closer look at A1 reveals that most of the flagged Beowulfian content is contained in fitt iv--the portion adjacent to B2. If A1 is divided in two--fitt iv and the remainder which we will call A1′--then the Beowulfian occurrence per line score decreases by 30% for A1′ while fitt iv scores the highest of all chunks investigated.

Table 3: Distribution of word compounds in A1 sub-divisions

	Beowulf parallels	Per line	*Beowulf* compounds	Per line	Cynewulf compounds	Per line	Unique compounds	Per line
A1	8	0.030	9	0.033	6	0.022	32	0.119
A1′	3	0.020	3	0.020	1	0.007	16	0.105
fitt iv	5	0.043	6	0.051	5	0.043	16	0.137
Lines 359-376	3	0.167	2	0.111	0	0.000	1	0.056

Half of the Beowulfian markers in fitt iv are found between lines 359 and 376 which also contains no Cynewulfian compounds:

> The holy one sat himself near the helmsman, noble one by noble one. I had never heard of a more beautiful ship ladened with splendid treasure. The warriors sat there--lords full of glory, beautiful thanes. Then spoke the mighty king, eternal Almighty, ordering his angel--glorious young retainer--to go and give meat to comfort the poor ones over flood's whelming, that they might more easily endure their condition over the waves' throng. Then it became disturbed, that shaken sea. The whale froliced, glided through the ocean, and the gray gull roamed eager to prey on the dead. The weather-candle faded, winds waxed, waves crashed, currents stirred, ropes snapped, and garments were soaked through.

Dreadful waters stood, pressing with force. The thanes became terrified.

(359-76a)[8]

This particularly Beowulfian passage embedded within A1 fills the first of the narrative

gaps observed in B above--Andrew's boarding of the boat and the beginning of the

storm. Without the A context, God's command to the attending angel in lines 364-6

need not be interpreted as a reference to the boat's crew. This supports the hypothesis

that the core poem did not include a divine helmsman, making the device a later

addition by the A poet.

One might argue that the sea voyage exhibits high affinity with *Beowulf* not

because of divergent textual history, but because the *Andreas* poet intentionally

referenced *Beowulf* during composition in light of the common content. This argument

does not well explain the high Beowulfian affinity in B3 which has fewer themes and

motifs in common with *Beowulf* nor the extremely low scores for A2-4. It would be

---

[8]The translation is my own. The original reads as follows with Beowulfian parallels underlined and compound words bolded.

>Gesæt him þa se halga   helmwearde neah,
>æðele be æðelum.   Æfre ic ne hyrde
>þon cymlicor   ceol gehladenne
>heah**gestreonum**.   Hæleð in sæton,
>þeodnas þrymfulle,   þegnas wlitige.
>Ða reordode   rice þeoden,
>ece ælmihtig,   heht his engel gan,
>mærne maguþegn,   ond mete syllan,
>frefran feasceafte   ofer flodes wylm,
>þæt hie þe eað mihton   ofer yða geþring
>drohtaþ adreogan.   Þa gedrefed wearð,
>onhrered hwælmere.   **Hornfisc** plegode,
>glad geond garsecg,   ond se græga mæw
>wælgifre wand.   Wedercandel swearc,
>windas weoxon,   wægas grundon,
>streamas styredon,   strengas gurron,
>wædo gewætte.   **Wæteregsa** stod
>þreata þryðum.   Þegnas wurdon
>acolmode.   (359-76a)

striking that so few Beowulfian parallels and formulae exist in A if the poet had access

to such a resource. Such an interpretation is further weakened if the A/B types can be

shown to have distinctive affinity for their own type, which would further support the

hypothesis of a composite nature. This possibility is investigated below using cluster

analysis.

Finally, there is the question of how Andrew was captured in the original B poem.

Like the passage discussed above from fitt iv, one would expect Beowulfian

characteristics to be correlated with those events in the A4 chunk. In A4, two short

passages are correlated with both Beowulfian parallels and compounds. The first

describes Andrew's capture:

Drogon deormodne   æfter dunscræfum,

ymb stanhleoðo,   stærcedferþne,

efne swa wide swa   wegas to lagon,

<u>enta **ærgeweorc**</u>,   innan burgum,

stræte **stanfage**.   (1232-6a)

[The brave-minded one was dragged by the cruel-hearted along mountain

caves, around rocky slopes, even so far as the sea-way, by the ancient

work of the giants, within the cities with stone-cobbled streets.]

The second passage begins Andrew's lament:

Næfre ic geferde   mid frean willan

<u>under heofonhwealfe   heardran</u> drohtnoð,

þær ic dryhtnes æ   deman sceolde.

Sint me leoðu tolocen,   lic sare gebrocen,

banhus **blodfag**,   benne weallað,

seonodolg swatige.   (1401-6a)

[Never have I born by the Lord's will such a sore living under the vault of

heaven where I must deem life the Lord's. My limbs are separated, my

body sorely broken, bone-house blood-stained, wounds welling up,

sinew-wounds sweaty.]

It is hard to say how much of the surrounding lines might go back to the original

core poem, but one can see how these lines might encourage expansion into a longer

passion set piece.

Evidence from the moving ratio of þ to ð

If the "core poem" hypothesis is accurate, then the A/B distinction should be

detectable via other methods sensitive to differences in textual history. Drout and

Chauvet have shown that by calculating the moving ratio of þ to ð in Anglo-Saxon

poetry and noting discrepancies in its variation--such as a change in trend

direction--one can generate "evidence of differences in textual history for particular

segments of Old English poems" as the orthographic tendencies of an exemplar leak

into the copied text (315-6).

Figure 4 plots the moving ratio of þ to ð for *Andreas* using a window of 1,000

words, overlayed with manuscript fitt demarcations and B section boundaries. Each B

section correlates to a change in trend of the ratio. The initial downtrend stops

temporarily at the end of B1. The end of B2 begins a new uptrend which reverses with the start of B3. Finally, the beginning of B4 correlates with the last uptrend in the ratio.
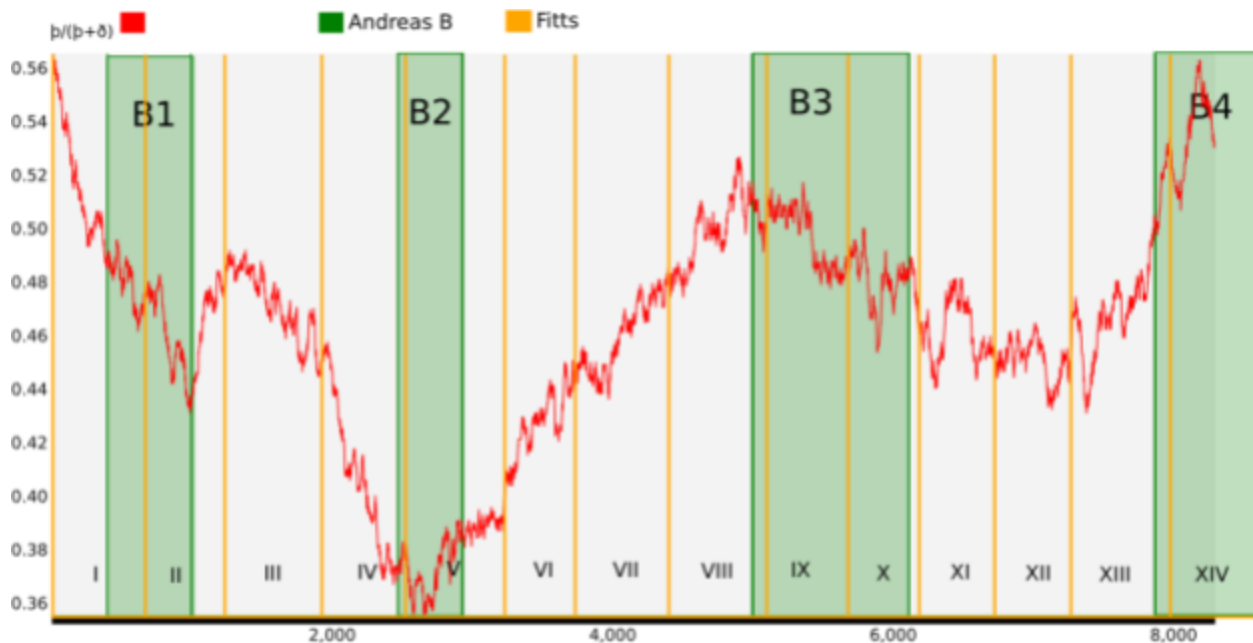
Figure 4: The moving ratio of þ to ð in *Andreas,* 1,000 word window

The size of the rolling window used to calculate a moving ratio affects the sensitivity of the indicator. Larger windows can filter noise from the time series, but can also obscure effects due to more complex textual structures ("Tracking" 291-2). A smaller window can be more responsive to changes in the feature being measured and therefore be able resolve smaller features, but this comes at the expense of additional noise as each instance carries greater statistical weight in smaller windows. It is good practice then to deploy larger windows for identifying major trends, then apply smaller windows to "zoom in" on areas of particular interest where the presence of smaller features is suspected.
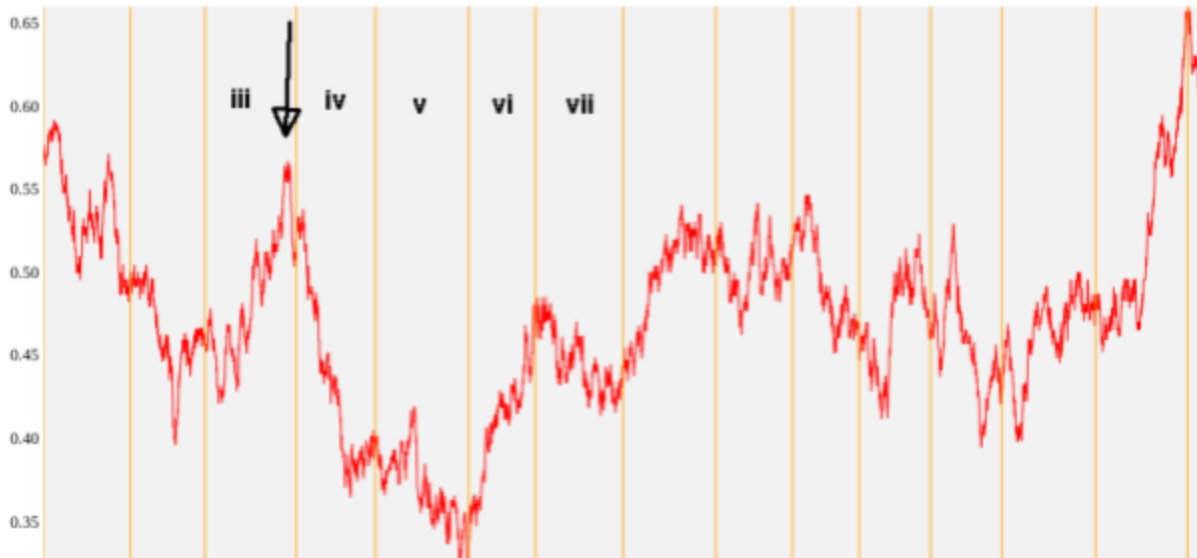
Figure 5: The moving ratio of þ to ð in *Andreas*, 500 word window

Calculating the ratio with a window of 500 words brings out a feature at the break between fitts iii and iv, correlating to the anomalous Beowulfian features in fitt iv described above (Figure 5). A change in trend between fitts vi and vii also marks a possible milestone for chunking this large A section for cluster analysis. Fitt vii marks the beginning of the story of the stone angel come to life.

Drout and Chauvet also tentatively hypothesize that a low level of þ (or equivalently, a higher percentage of ð) could indicate older provenance as ð may have entered Old English orthography first (295). Three of the four B sections occur during downtrends in the ratio, which indicates higher ð counts. The sea voyage passage in particular records an extremely high ð percentage.

This moving ratio quantifies orthographic tendencies, not stylistic language. Therefore, it can be considered an independent line of evidence in favor of the A/B hypothesis. Passages with Beowulf affinity are correlated with a decrease in the

proportion of þ which suggests that they may be relatively older than passages with Cynewulfian affinity.

Evidence for the A/B division from cluster analysis

Having established a quantitative basis for the A/B division in *Andreas*, hierarchical, agglomerative cluster analysis can be used to establish whether the sections of each type show greater affinity for their own type. This technique has been used to detect internal structure in *Guthlac A* and its Latin exemplar, *Genesis A and B*, *Beowulf,* and the Christ poems (Downey; *Unlocked*; "Dendrogrammatology").

Manuscript fitt and B section boundaries were used to divide the poem into ten chunks. Table 4 lists the chunks and word count contained in each. Experimental evidence has shown that cluster analysis is efficacious when chunks are between 400 and 1,500 words ("Dendrogrammatology" 313). Sections showing affinity to *Beowulf* have been subdivided in order to facilitate possible blending strategies described below.

Table 4: Chunk sizes by line count

intro	B1	A1 & B2			A2 & A3		B3	A4	B4	end
		A1′	fitt iv	B2	A2	A3				
400	597	920	560	484	762	1284	1100	1761	852	568
		1480		1044	2046					

Performing a basic cluster analysis excluding the introduction and epilogue, B sections tend to cluster separately from A sections, with the shortest A chunk as the exception (Figure 6). The shortest chunk B2 is least alike of all, clustering as a

simplicifolious clade. In fact, the stepwise nature of the dendrogram suggests that

several chunks may be too small to quantify their word frequency profiles sufficiently.
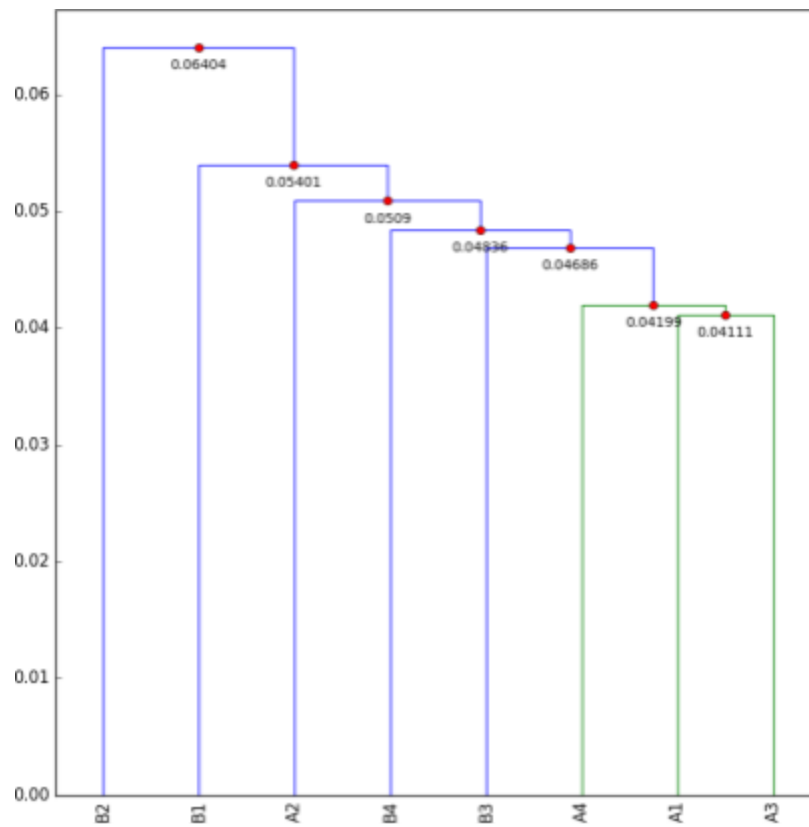
Figure 6: Dendrogram of unblended A/B chunks

In *Beowulf Unlocked*, Drout suggests blending as one possible strategy for

dealing with subsections which are too short or interleaved (12). Several possible

blending strategies are evident from the analysis thus far:

1. Blend B1 and B2 as both are short B sections.

2. Blend fitt iv from A1 with B2 in light of its *Beowulf* affinity.

3. Blend A2 and B2 as both contain conversations with Jesus.

4. Blend A2 and A3 because the þ/ð and Cynewulfian formulaic trends are

   continuous.

5. Blend the introduction with B1 and B4 with the epilogue because formulae
   are sparse on the edges.

6. Blend the introduction and epilogue.

The techniques of incrementation and truncation[9] were used to observe how each blending strategy affected the dendrogram (*Unlocked* 13). Across blending strategies, A sections tend to cluster together in a clade while B either clusters on its own or as a stepwise pattern separate from the A clade. When B2 is blended with either B1 or A2, the resulting chunk will cluster within the A clade. If B2 is blended with fitt iv, then it pulls A1′ out of the A clade and they combine to create a clade of especially Beowulfian material (Figure 7). Note also that the position of the Beowulfian clade shows more affinity to B than A.

---

[9] Incrementation builds up a dendrogram one chunk at a time while truncation is the opposite. This allows the critic to isolate how a single chunk affects the dendrogram geometry.
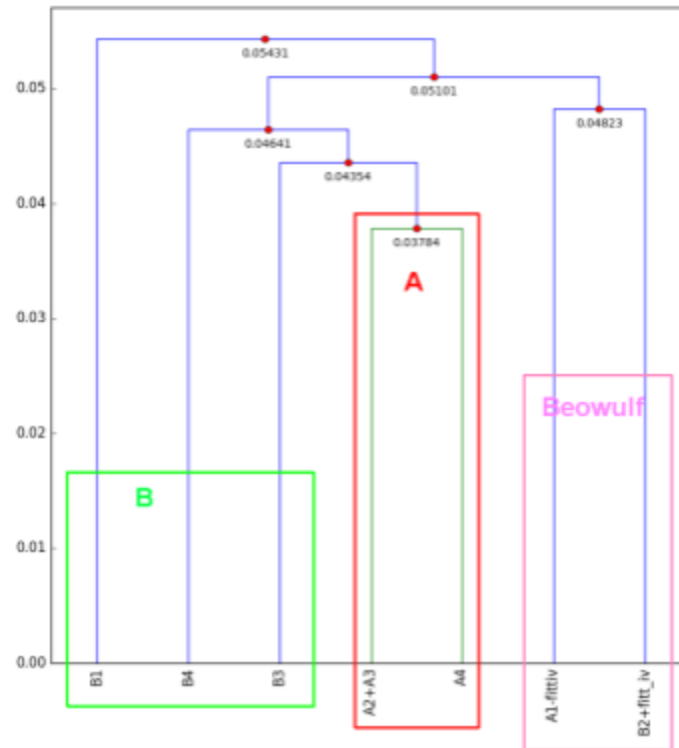
Figure 7: Dendrogram of blended sections

When the edge chunks are introduced either on their own, blended together, or blended with their adjacent B sections, the edge chunks tend to show an affinity with B. However, when the epilogue is blended with B4 it will cluster in the A clade suggesting that the epilogue has more affinity with A than the introduction does (Figure 8). It is worth noting that although the edge chunks have the same number of Cynewulfian formulae, they are more evenly distributed in the epilogue; the formulae in the introduction are bunched up near the B1 boundary.
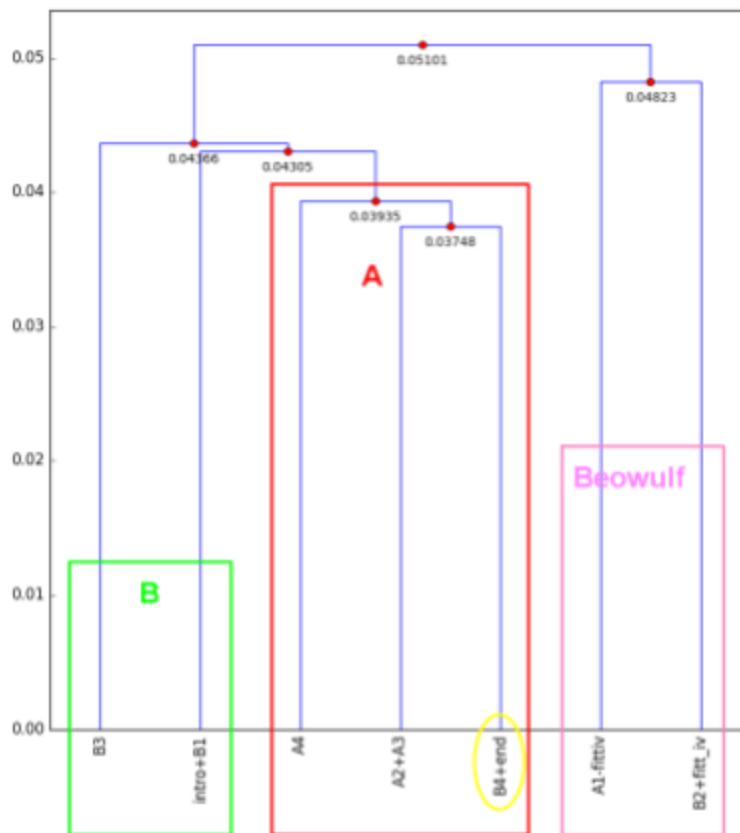
Figure 8: Dendrogram with edge sections

Cluster analysis supports the view that A and B sections have an affinity within themselves and that there exists a Beowulfian affinity between A1 and B2. It also shows that although the edge chunks have a general affinity to B, the end exhibits relatively more affinity to A than the beginning.

Cluster analysis within the Cynewulfian group

Thus far, cluster analysis has been deployed for the purpose of investigating the internal structure of the poem, but it can also be used to compare relative affinity within a set of texts ("Dendrogrammatology" 323-5). When doing so, one must be careful to

account for possible orthographic differences such as spelling variation or þ/ð

tendencies of different scribes which could bias the results (*Unlocked* 17-8). For the

following cluster analysis, common variances were consolidated and the top fifty most

frequent words were manually checked for spelling variation.[10]

The A and B sections were injected into a dendrogram made up of the wider

Cynewulfian group of poems as described by Drout et al in "Of Dendrogrammatology"

(325).[11] Signed poems by Cynewulf form their own clade[12] along with *Guthlac B*, which

has a strong claim to Cynewulfian authorship itself (Fulk 5). If both sets A and B are

injected at the same time all of *Andreas* will collapse into its own clade, but taken

separately one can see how they each have a different level of affinity for the signed

Cynewulf works. *Andreas A* clusters within the innermost Cynewulfian clade along with

the majority of *Elene* (Figure 9). However, when *Andreas B* is injected only B3 clusters

within the *Elene* sub-clade while the rest is relegated to the fringe of the larger

Cynewulfian clade (Figure 10).

---

[10] *And*, *ond*, and tironian note were consolidated together as were all thorn and eth characters.
[11] The naming convention used is E-- *Elene*, J--*Juliana*, F--*The Fates of the Apostles*, CII--*Christ II*, GA--*Guthlac A*, GB--*Guthlac B*. Chunks are roughly 1,000 words each with boundaries adjusted for known divisions.
[12]Most of *Juliana* does not cluster within the Cynewulfian clade. Drout speculates that this is due to particular influence from its Latin source (333).
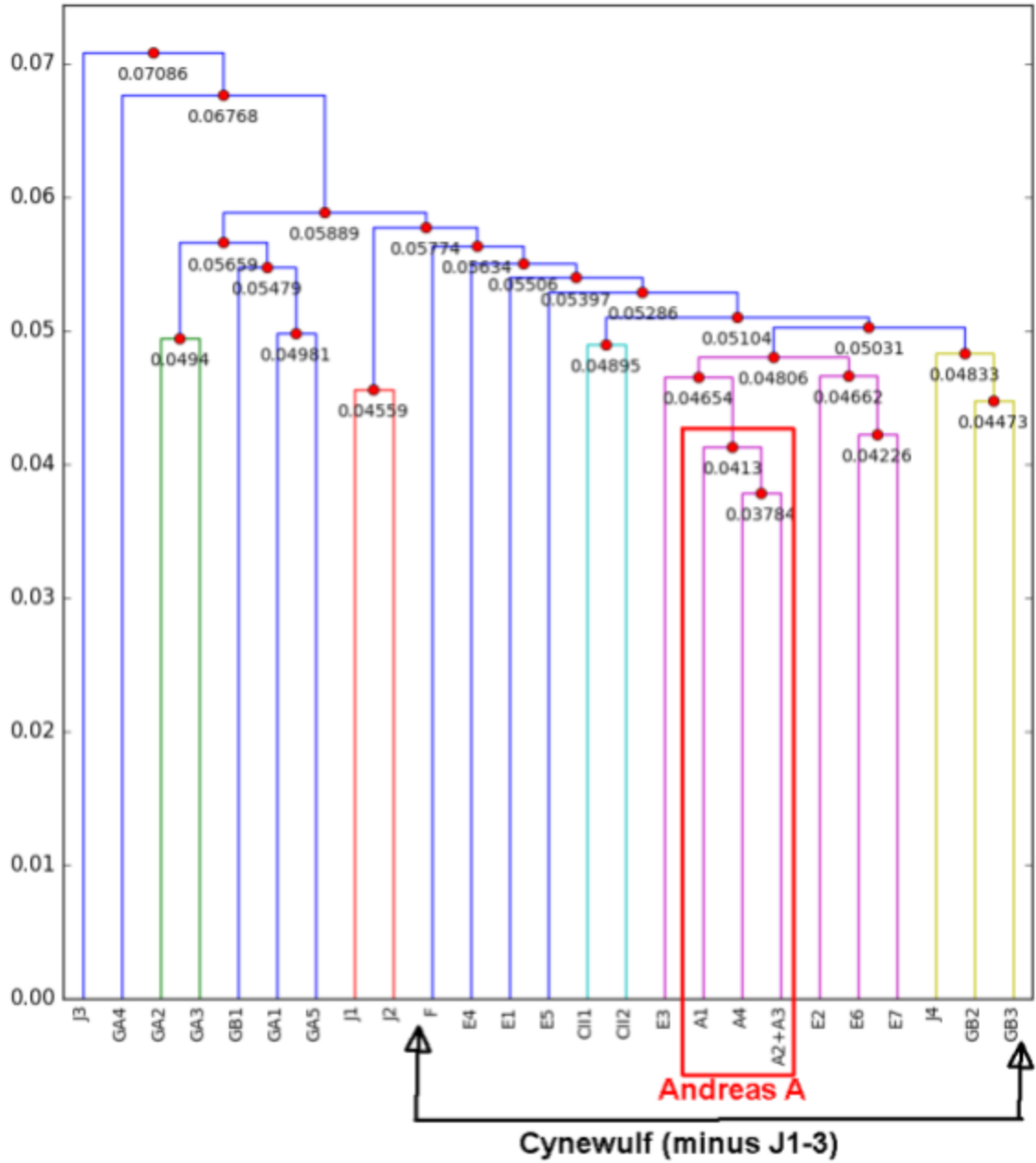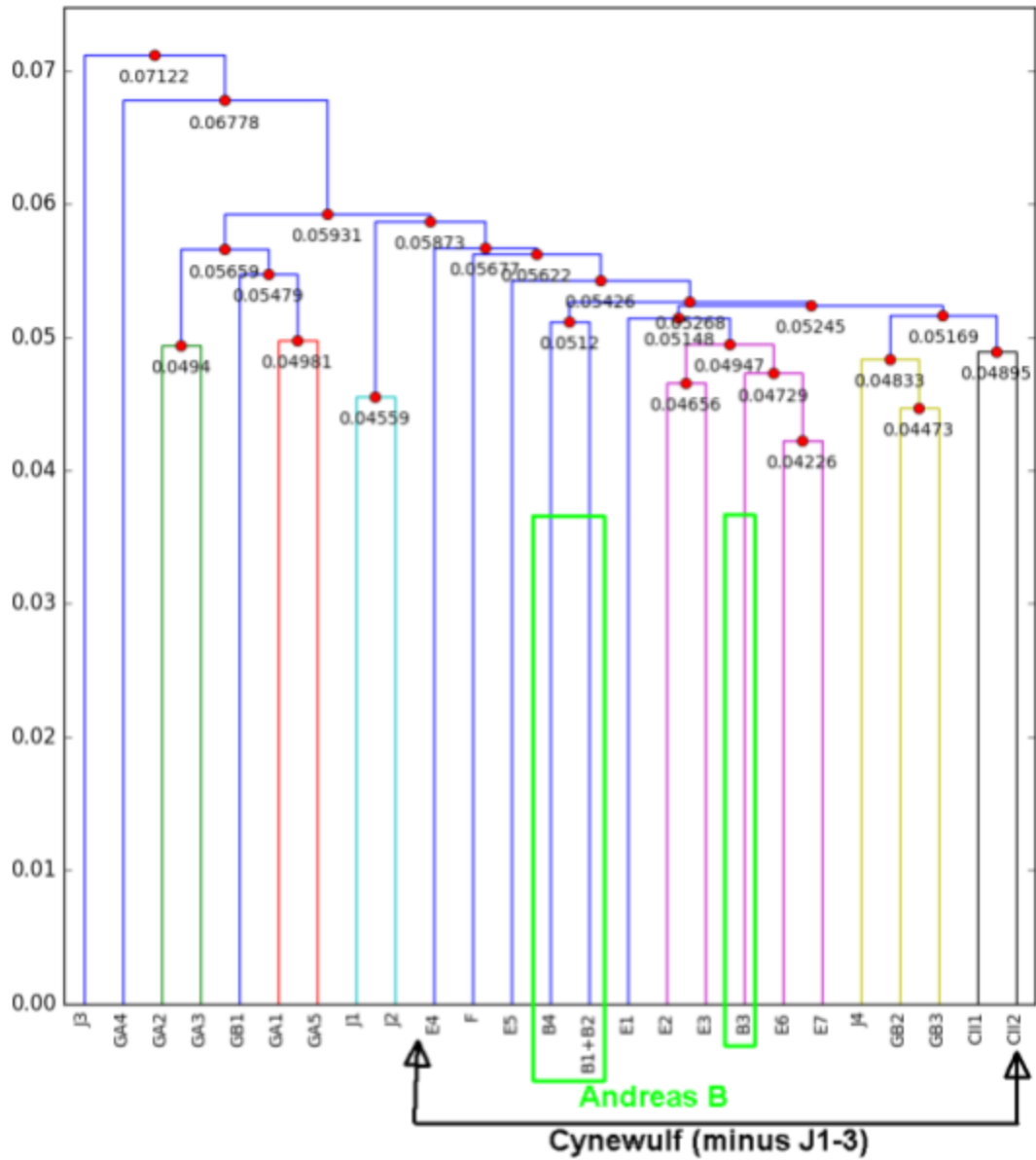
Figure 9: *Andreas A* with the Cynewulfian group

Figure 10: *Andreas B* with the Cynewulfian group

Cluster analysis across the Cynewulfian group shows that the A sections have a

stronger affinity with signed Cynewulf work than the B sections. The clustering of B3

within the *Elene* sub-clade accords with the evidence from the distribution of

Cynewulfian word compounds, further supporting the idea that this section has been reworked more by the A poet.

Arguments against Cynewulfian authorship of *Andreas*

Several critics have noted grammatical and stylistic features that distinguish *Andreas* from the signed poems.[13] George Krapp outlined such a list of differences to which R.D. Fulk added several metrical distinctions (Krapp xlviii-xlix; Fulk 7-8). These items can be divided into two types:

1. Features rare or non-existent in Cynewulf but that appear in *Andreas*, or

2. Features often found in Cynewulf that are absent from or rare in *Andreas*.

By definition, the latter items are not quantifiable since they do not appear in *Andreas*. However, "non-Cynewulfian" words in *Andreas* can be correlated to the A/B divisions previously described. Fulk cites eight such differences taken from Krapp: *ondswarode*, dative *fæder*, *sin*, *æninga*, *becweðan*, *feorr*, *æfter þam/þyssum wordum*, and *wordum/worde cwæð*. These represent fifty-one instances. He also argues that, unlike the signed poems, *Andreas* does not vary metrical types used when resolving certain kinds of compound words (8). Almost 75% of these phrases are in the A sections despite the fact that the A lines constitute only 57% of the poem.

---

[13]The question of whether these discrepancies constitute evidence of distinct authorship or different periods of activity or context for the same author bears further inquiry. Most cases enumerated by Krapp and Fulk are statistical arguments, not clear absolute distinctions. While working with his Delta procedure, Burrows noted that the "uncharacteristic" work of known authors can yield inaccurate statistical results, such as changes in style over the span of a long career or substantial change in genre or subject matter ("Delta" 279; "Questions" 21-3).

Table 5: The distribution of Fulkian anomalies

	Intro	B	A	End
# of Fulk items	2	11	38	0
% of Fulk items	3.92%	21.57%	74.51%	0%
% of lines	4.36%	32.87%	56.68%	6.10%

Most of the additional differences listed by Krapp also predominantly appear in the A sections, including the phrases *ða gen/git*, *sum* + a genitive plural, *eft swa ær, lyt,* and an overuse of *siððan*. As the anomalies enumerated by Krapp and Fulk disproportionately apply to the A sections of the poem, this line of inquiry also supports the A/B hypothesis.

Finally, Krapp concludes his list of anomalies by citing Fritzsche's observation that, unlike signed poems, *Andreas* "nowhere alludes to any written sources" (xlix). However, in lines 1478-91 the poet breaks into the narrative by referring to himself in the first person. He expresses his humility as a poet and reminds the reader of the *fyrn-sægen* (ancient sayings) about Andrew's story. Although this passage is contained in the B4 chunk, this may be an interpolation by the A poet referring to the original core itself--a reference which may have been obvious to a contemporary audience familiar with the earlier B poem. This passage has no parallels to *Beowulf* as identified by Orchard but does contain a Cynewulfian compound word and two compounds unique to *Andreas*, of which *fyrn-sægen* is one example.[14] Downey et al found a similar practice in effect in the Guthlac poems. *Guthlac A* and *B* each contain references to *bec* (books)

---

[14]*Leoð-word* (poetic word) is the other. The resulting occurrence per line scores are 0.071 Cynewulfian, 0.143 Andrean. Both score greater than 20% above the mean for their respective metrics. *The Dictionary of Old English* confirms that this is the only instance of *fyrn-sægen* in poetry.

correlating with detectable changes in source (23). Although one may grant that *bec* has

a stronger implication of "written" sources than *fyrn-sægen*, the passage fills the same

role as those found in *Guthlac* and the signed poems. Furthermore, if the A poet saw

the B poem as divergent from his written Latin sources, then *fyrn-sægen* may have

been a more accurate description of the source from his point of view.

Conclusion

Modern quantitative techniques should not be wielded uncritically or in isolation,

but combined with traditional qualitative analysis they offer a new avenue by which to

approach well-known texts. Quantitative data can also serve as a neutral arbiter

between incompatible qualitative assessments. *Andreas* has been interpreted widely by

different scholars as an early poem with *Beowulf*-like characteristics, an unsigned

product of Cynewulf, and a post-Cynewulfian homage.

Although it may be possible to explain the evidence of an A/B division in *Andreas*

as the product of textual influence from different reference texts on a singular poet, the

conclusion that *Andreas* consists of an earlier core which has been expanded upon is

more parsimonious with the quantitative data. Analysis of oral-formulaic language, word

frequency and orthographic tendencies point toward a new understanding of *Andreas*

as a composite work which is further supported by a close reading of the text.

Therefore, the textual history of *Andreas* can be reconstructed tentatively as follows.

An early version of the poem still detectable as *Andreas B* was composed first,

probably by a pre-Cynewulf poet. It relied heavily on heroic diction in the tradition of

Beowulf and was more concerned with telling an heroic tale than teaching Christian doctrine. That certain features found in Latin sources are not present in B suggests that this poem represents an earlier form of the story, or at least one with a particularly Anglo-Saxon flavor. In addition to the B sections, it likely included much of the introduction and fitt iv, especially lines 359-76. Remnants may also survive in A4 and the ending passage.

Taking this poem as a foundation, the A poet expanded the B poem with material identified above as *Andreas A* including overtly didactic, moralizing episodes such as the divine helmsman, the discursive stories of sceptical unbelievers, and Andrew's *imitatio Christi*. Harmonizing the B poem with a known Latin source may have been a primary motivation for the A revisions. The B3 section and ending passage appear to have been more heavily edited by the A poet than the first half of B. However, a favorite stylistic tic of the A poet was coining novel compound words which he used throughout the B sections, even in areas where he preferred to retain Beowulfian phraseology. The A poet exhibits striking affinity with the signed poems which suggests that he was either Cynewulf himself or someone very well acquainted with his work.

For *Andreas B*, much work remains to explore how it fits into the history of Old English poetry, who the author may have been, where and when it may have been written, and what new information it may provide us when considered as a newly re-discovered poem in its own right. In the case of *Andreas A*, the question of Cynewulfian authorship needs to be re-evaluated in light of the implications of the A/B

division on the style of the A poet, both as the author of the A sections and as editor of the B material.

Early critics of Old English poetry, and the Cynewulfian group in particular, were characterized by an eagerness to ascribe all manner of unsigned poems to a few known poets such as Cædmon and Cynewulf. As later scholars adopted more conservative interpretations, the pendulum of scholarly consensus swung toward a minimal Cynewulfian canon. Quantitative analysis offers us a way to temper the swings--to anchor discourse to more objective metrics and define the bounds wherein subjective interpretation applies. The composite nature of *Andreas* as identified above must be seen as one small nudge of the pendulum toward a more inclusive interpretation of Cynewulfian authorship.

Works Cited

Anderson, Earl R. "The Structure of *Elene*." C*ynewulf: Structure, Style and Theme in His Poetry.* Fairleigh Dickinson UP, 1983, pp. 103-25.

Bjork, Robert E. "Introduction." *The Cynewulf Reader*, edited by Robert E. Bjork, Routledge, 2001, pp. xv-xxiii.

Burrows, John. "'Delta': a Measure of Stylistic Difference and a Guide to Likely Authorship." *Literary and Linguistic Computing*, vol. 17, no. 3, Jan. 2002, pp. 267–287.

Burrows, John. "Questions of Authorship: Attribution and Beyond: A Lecture Delivered on the Occasion of the Roberto Busa Award ACH-ALLC 2001, New York." *Computers and the Humanities*, vol. 37, no. 1, 2003, pp. 5–32.

Downey, Sarah, et al. "'Books Tell Us': Lexomic and Traditional Evidence for the Sources of *Guthlac A*." *Modern Philology*, vol. 110, 2012, pp. 1–29.

Drout, Michael D.C., et al. *Beowulf Unlocked: New Evidence from Lexomic Analysis.* Springer International Publishing, 2016.

Drout, Michael D. C., et al. "Of Dendrogrammatology: Lexomic Methods for Analyzing Relationships among Old English Poems." *The Journal of English and Germanic Philology*, vol. 110, no. 3, 2011, pp. 301–336.

Drout, Michael D. C., and Elie Chauvet. "Tracking the Moving Ratio of þ to ð in Anglo-Saxon Texts: A New Method, and Evidence for a Lost Old English Version of the 'Song of the Three Youths'". *Anglia*, vol. 133, 2015, pp. 278-319.

Friesen, Bill. "Visions and Revisions: The Sources and Analogues of the Old English *Andreas*." Diss. U of Toronto, 2009.

Fulk, R. D. "Cynewulf: Canon, Dialect, and Date." *The Cynewulf Reader*, edited by Robert E. Bjork, Routledge, 2001, pp. 3–21.

Greenfield, Stanley B., et al. *A New Critical History of Old English Literature*. New York University Press, 1986.

Hoover, D. L. "Testing Burrows's Delta." *Literary and Linguistic Computing*, vol. 19, no. 4, Jan. 2004, pp. 453–475.

Krapp, George Philip. *Andreas and The Fates of the Apostles*. Boston, 1905, pp. xlviii-xlix.

Moretti, Franco. *Graphs, Maps, Trees: Abstract Models for Literary History*. Verso, 2007.

Neidorf, Leonard. "The Composite Authorship of *The Dream of the Rood*." *Anglo-Saxon England 45,* pp. 51–70. Web.

Orchard, Andy. "Both Style and Substance: The Case for Cynewulf." *Anglo-Saxon Styles*, edited by Catherine Karkov, State University of New York Press, 2003, pp. 294–296.

Orchard, Andy. "The Originality of *Andreas*." *Old English Philology: Studies in Honour of R.D. Fulk.* By Leonard Neidorf, Rafael J. Pascual, and T. A. Shippey. Cambridge, UK: D. S. Brewer, 2016. 331-70. Print.

Pierce, John R. *An Introduction to Information Theory: Symbols, Signals & Noise*. Dover Publications, 1997.

Stevick, Robert D. "Arithmetical Design of the Old English *Andreas*." *Anglo-Saxon Poetry: Essays in Appreciation, for John C. McGalliard*, by John C. McGalliard et al., University of Notre Dame Press, 1975, pp. 99–115.